On L^p-strong convergence of an averaging principle for non-Lipschitz slow-fast systems with Lévy noise

Yong Xu[a,b,*], Hongge Yue[a], Jiang-Lun Wu[c]

[a]*Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, 710072, China*
[b]*MIIT Key Laboratory of Dynamics and Control of Complex Systems*
[c]*Department of Mathematics, Computational Foundry, Swansea University, Swansea SA1 8EN, UK*

Abstract

We study L^p-strong convergence for coupled stochastic differential equations (SDEs) driven by Lévy noise with non-Lipschitz coefficients. Utilizing Khasminkii's time discretization technique, the Kunita's first inequality and Bihari's inequality, we show that the slow solution processes converge strongly in L^p to the solution of the corresponding averaged equation.

Keywords. Slow-fast systems, averaging principle, non-Lipschitz coefficients, Lévy noise.

Mathematics subject classification. 70K70, 60H10, 34K33

1. Introduction

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space equipped with a natural filtration $\{\mathscr{F}_t, t \geq 0\}$, and $\|\cdot\|$ denote the Euclidean norm as well as the matrix trace norm. Consider the following slow-fast stochastic differential equations with Lévy noise

$$\begin{cases} dX_t^\epsilon = a\left(X_t^\epsilon, Y_t^\epsilon\right) dt + b\left(X_t^\epsilon\right) dB_t^1 + \int_{\mathbb{Z}} \sigma\left(X_t^\epsilon, z\right) \tilde{N}_1\left(dt, dz\right), X_0^\epsilon = x_0, \\ dY_t^\epsilon = \frac{1}{\epsilon} f\left(X_t^\epsilon, Y_t^\epsilon\right) dt + \frac{1}{\sqrt{\epsilon}} g\left(X_t^\epsilon, Y_t^\epsilon\right) dB_t^2 + \int_{\mathbb{Z}} h\left(X_t^\epsilon, Y_t^\epsilon, z\right) \tilde{N}_2^\epsilon\left(dt, dz\right), Y_0^\epsilon = y_0, \end{cases} \tag{1.1}$$

for small parameter $0 < \epsilon \ll 1$ and certain mappings $a(u,v) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, b(u) : \mathbb{R}^n \to \mathbb{R}^{n \times d_1}, \sigma(u,z) : \mathbb{R}^n \times \mathbb{Z} \to \mathbb{R}^n, f(u,v) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m, g(u,v) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{m \times d_2}, h(u,v,z) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{Z} \to \mathbb{R}^m$. The driving processes B_t^1 and B_t^2 are respectively d_1- and d_2-dimensional independent Brownian motions, $\tilde{N}_1(dt,dz) = N_1(dt,dz) - \nu_1(dz)\,dt$, $\tilde{N}_2^\epsilon(dt,dz) = N_2(dt,dz) - \frac{1}{\epsilon}\nu_2(dz)\,dt$ are compensated Poisson random measures, respectively, with intensity Lévy measure ν_1 and ν_2 on a measurable space $(\mathbb{Z}, \mathscr{B}(\mathbb{Z}))$ of $(-\infty, +\infty)$, where $\mathscr{B}(\mathbb{Z})$ denotes the Borel σ-algebra of $\mathbb{Z}$. Recently, [1] proposed a set of conditions on the non-Lipschitz coefficients (which is more general and much weaker than the usual Lipschitz conditions) ensuring the existence and uniqueness of solutions to the system (1.1).

The theory of averaging principle for SDEs was initiated by Khasminkii [2]. Since then, there are vast development devoted to this topic. For the slow-fast systems with noise, the

*Corresponding author
Email addresses: hsux3@nwpu.edu.cn (Yong Xu), yuehongge803@163.com (Hongge Yue), j.l.wu@swansea.ac.uk (Jiang-Lun Wu)

averaging principle provides a powerful tool for approximating the slow-fast SDEs by reduced SDEs with slow component, see the recent work for systems with Gaussian white noise in [3, 4], the systems with Lévy-type perturbations in [5–7] and with non-Lipschitz coefficients in [8–10]. While most of the above mentioned references dealt mainly with square convergence, it is a natural and also very important question: Whether the validity of averaging principle to non-Lipschitz slow-fast systems still holds under higher order-convergence? Our aim in the present paper is to establish the convergence scheme in L^p sense (i.e., convergence in p-th moments) for the slow-fast system (1.1).

The rest of the paper is organized as follows. Section 2 presents some assumptions and preliminary results that are needed in the subsequent section. In Section 3, we prove that the slow component processes X_t^ϵ converge to the limit process $\bar{X}_t$ in the sense of p-th moments. Throughout the paper, C, C_p, C_T and $C_{p,T}$ denote positive constants which may change from line to line, where C_p depends on p, C_T depends on T, and $C_{p,T}$ depends on p, T, etc..

2. Preliminaries

We assume the coefficients of system (1.1) fulfil the following

(A1) There exists a concave continuous nondecreasing function $\kappa(\cdot)$, such that

$$\begin{aligned}\|a(u_1,v_1)-a(u_2,v_2)\|^q+\|b(u_1)-b(u_2)\|^q+\int_{\mathbb{Z}}\|\sigma(u_1,z)-\sigma(u_2,z)|^q\nu_1(dz)\\ \le C\kappa\left(\|u_1-u_2\|^q+\|v_1-v_2\|^q\right),\end{aligned}$$

$$\begin{aligned}&\|f(u_1,v_1)-f(u_2,v_2)\|^q+\|g(u_1,v_1)-g(u_2,v_2)\|^q\\ &\qquad+\int_{\mathbb{Z}}\|h(u_1,v_1,z)-h(u_2,v_2,z)\|^q\nu_2(dz)\\ &\qquad\qquad\le C\left(\|u_1-u_2|^q+\|v_1-v_2\|^q\right),\end{aligned}$$

for all $u_j\in\mathbb{R}^n, v_j\in\mathbb{R}^m (j=1,2)$, $q\ge 2$, where $\kappa(\cdot):\mathbb{R}^+\to\mathbb{R}^+$, $\kappa(0)=0$, $\int_{0^+}\frac{du}{\kappa(u)}=\infty$. One can find a pair of positive constants α_1 and α_2 such that for $\kappa(u)\le\alpha_1 u+\alpha_2, \forall u\ge 0$.

(A2) Assume that a is globally bounded, a, b, σ, f, g and h satisfy the linear growth conditions. Precisely, there exist constants ι_1, ι_2 and ι_3 such that

$$\begin{aligned}&\|a(u,v)\|^2\vee\|b(u)\|^2\vee\|f(u,v)\|^2\vee\|g(u,v)\|^2\le\iota_1(1+\|u\|^2+\|v\|^2),\\ &\int_{\mathbb{Z}}\|\sigma(u,z)\|^q\nu_1(dz)\le\iota_2(1+\|u\|^q),\quad\int_{\mathbb{Z}}\|h(u,v,z)\|^q\nu_2(dz)\le\iota_3(1+\|u\|^q+\|v\|^q),\end{aligned}$$

for all $u\in\mathbb{R}^n, v\in\mathbb{R}^m, q\ge 2$.

(A3) There exist constants $c_0>0$ and $c_j\in\mathbb{R}$, $j=1,2$, which are all independent of (u_1,v_1,v_2), such that

$$v_1^{\mathrm{T}}f(u_1,v_1)\le -c_0\|v_1\|^2+c_1,\quad (f(u_1,v_1)-f(u_1,v_2))^{\mathrm{T}}(v_1-v_2)\le c_2\|v_1-v_2\|^2,$$

for all $u_1\in\mathbb{R}^n, v_1,v_2\in\mathbb{R}^m$, where T stands for the transpose.

(A4) There exists a constant $\beta > 0$, which is independent of (u, v), such that

$$v^T g(u,v) g^{\mathrm{T}}(u,v) v \geq \beta \|v\|^2.$$

Remark 2.1. According to Assumptions (A1)-(A2), the existence and uniqueness of solutions of system (1.1) can be obtained (see [1]). On the other hand, Assumptions (A3)-(A4) are two conditions which yield a unique invariant measure possessing exponentially mixing property to the fast variable motion (see [11]).

Now, we define the solution of the corresponding averaged equation:

$$d\bar{X}_t = \bar{a}(\bar{X}_t)dt + b(\bar{X}_t)dB_t + \int_{\mathbb{Z}} \sigma(\bar{X}_t, z)\tilde{N}_1(dt, dz),\ \ \bar{X}_0 = x_0, \tag{2.1}$$

with the averaged coefficient $\bar{a}(x) = \int_{\mathbb{R}^m} a(x,y)\mu^x(dy)$, and μ^x is the unique invariant measure for the transition semigroup of the frozen equation

$$dY_t = f\left(x, Y_t\right) dt + g\left(x, Y_t\right) d\bar{B}_t + \int_{\mathbb{Z}} h\left(x, Y_t, z\right) \bar{N}\left(dt, dz\right), Y_0 = y_0. \tag{2.2}$$

for any fixed $x \in \mathbb{R}^n$.

Remark 2.2. For $x \in \mathbb{R}^n, y \in \mathbb{R}^m$, one has $\|\mathbb{E}a(x, Y_t^{x,y}) - \bar{a}(x)\|^2 \leq Ce^{-\eta t}(1 + \|x\|^2 + \|y\|^2)$, where η is a positive constant concerned with $\iota_i (i = 1, 3)$ and $c_j (j = 0, 1, 2)$. This is due to the invariant property of μ^x (see [12]).

Lemma 2.3. *Assume that* (A1)-(A4) *are satisfied, then there exists a positive constant* $C_{p,\iota_i,c_j,T}$ *such that for all* $\epsilon \in (0,1), t \in [0,T]$, $p \geq 1$,

$$\begin{aligned}
\mathbb{E} \sup_{0 \leq t \leq T_0} \|X_t^\epsilon\|^{2p} &\leq C_{p,\iota_i,T}, \\
\sup_{0 \leq t \leq T_0} \mathbb{E}\|Y_t^\epsilon\|^{2p} &\leq C_{p,\iota_i,c_j,T},
\end{aligned}$$

where $C_{p,\iota_i,c_j,T}$ *is independent of* ϵ, $i = 1, 2, 3, j = 0, 1$.

Proof: The techniques similar to those used in Lemma 3.1 and 3.2 in [13]. Here, the detailed proof is omitted for the sake of brevity. □

Lemma 2.4. *Assume that* (A1)-(A4) *are satisfied, for all* $t \in [0,T], h \in (0,1)$, $p \geq 1$, *then there exists a positive constant* $C_{p,\iota_i,c_j,T}$ *such that*

$$\mathbb{E}\|X_{t+h}^\epsilon - X_t^\epsilon\|^{2p} \leq C_{p,\iota_i,c_j,T} h,$$

where $C_{p,\iota_i,c_j,T}$ *is independent of* ϵ, $i = 1, 2, 3,\ j = 0, 1$.

Proof: Thanks to the Hölder inequality, Assumption (A2), the Burkholder's inequality and Burkholder-Davis-Gundy inequality, the following can be derived

$$
\begin{aligned}
&\mathbb{E}\|X^{\epsilon}_{t+h}-X^{\epsilon}_{t}\|^{2p}\\
\leq\ & C_p\mathbb{E}\left\|\int_t^{t+h} a(X^{\epsilon}_s,Y^{\epsilon}_s)ds\right\|^{2p}+C_p\mathbb{E}\left\|\int_t^{t+h} b(X^{\epsilon}_s)dB^1_s\right\|^{2p}+C_p\mathbb{E}\left\|\int_t^{t+h}\int_{\mathbb{Z}}\sigma(X^{\epsilon}_s,z)\tilde{N}_1(ds,dz)\right\|^{2p}\\
\leq\ & C_p h^{2p-1}\mathbb{E}\int_t^{t+h}\|a(X^{\epsilon}_s,Y^{\epsilon}_s)\|^{2p}ds+C_p\mathbb{E}\left[\int_t^{t+h}\left\|b(X^{\epsilon}_s)\right\|^2 ds\right]^p\\
&+C_p\mathbb{E}\left[\int_t^{t+h}\int_{\mathbb{Z}}\left\|\sigma(X^{\epsilon}_s,z)\right\|^2\nu_1(dz)ds\right]^p+C_p\mathbb{E}\left[\int_t^{t+h}\int_{\mathbb{Z}}\left\|\sigma(X^{\epsilon}_s,z)\right\|^{2p}\nu_1(dz)ds\right]\\
\leq\ & C_{p,\iota_1}h^{2p-1}\int_t^{t+h}\mathbb{E}\left(1+\|X^{\epsilon}_s\|^{2p}+\|Y^{\epsilon}_s\|^{2p}\right)ds+C_{p,\iota_1}h^{p-1}\int_t^{t+h}\mathbb{E}\big(1+\|X^{\epsilon}_s\|^{2p}\big)ds\\
&+C_{p\iota_2}h^{p-1}\mathbb{E}\left[\int_t^{t+h}(1+\|X^{\epsilon}_s\|^2)^p ds\right]+C_{p\iota_2}\mathbb{E}\left[\int_t^{t+h}(1+\|X^{\epsilon}_s\|^{2p})ds\right]\\
\leq\ & C_{p,\iota_i,c_j,T}h.
\end{aligned}
$$

This completes the proof of Lemma 2.4. □

Remark 2.5. We point out that the slow compolent contains the compensated Poisson random measure in Lemma 2.4. This makes an intrinsic difference between Gaussian noise and Lévy-type noise to regularity of X^{ϵ}_t, which affecting the rate of convergence scale in practical engineering.

3. Stochastic averaging principle

In this part, we intend to prove that the slow component process X^{ϵ}_t convergence strongly to the solution $\bar{X}_t$ of the averaged equation.

Theorem 3.1. *Assume that* (A1)-(A4) *are satisfied, then if we choose* $\delta=\epsilon ln\epsilon^{-k}$*, for any* $T>0$, $p\geq 1$, *we obtain*

$$
\lim_{\epsilon\to 0}\mathbb{E}\sup_{0\leq t\leq T}\|X^{\epsilon}_t-\bar{X}_t\|^{2p}=0. \tag{3.1}
$$

Proof. The proof is divided into three steps. In Step 1, $\mathbb{E}\sup_{0\leq t\leq T}\left\|X^{\epsilon}_s-\hat{X}^{\epsilon}_s\right\|^{2p}$ will be estimated, we then derive estimate of $\mathbb{E}\sup_{0\leq t\leq T}\left\|\hat{X}^{\epsilon}_t-\bar{X}_t\right\|^{2p}$ in Step 2. At last, through Step 1 and Step 2, the main result of Theorem 3.1 will be obtained.

Step 1. Following the idea inspired by Khasminskii [2], we introduce an auxiliary processes $(\hat{X}^{\epsilon}_t,\hat{Y}^{\epsilon}_t)\in\mathbb{R}^n\times\mathbb{R}^m$. We consider a partition of $[0,T]$ into intervals of the identical length δ, and the time interval $\delta=\epsilon ln\epsilon^{-k}(k>0)$ is selected small enough. We construct a process $\hat{Y}^{\epsilon}_t$, with initial datum $\hat{Y}^{\epsilon}_t=y_0$,

$$
\hat{Y}^{\epsilon}_t=Y^{\epsilon}_{k\delta}+\frac{1}{\epsilon}\int_{k\delta}^t f(X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s)ds+\frac{1}{\sqrt{\epsilon}}\int_{k\delta}^t g(X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s)dB^2_s+\int_{k\delta}^t\int_{\mathbb{Z}}h(X^{\epsilon}_{k\delta},\hat{Y}^{\epsilon}_s,z)\tilde{N}^{\epsilon}_2(ds,dz),
$$

where $X_{k\delta}^{\epsilon}$ and $Y_{k\delta}^{\epsilon}$ are slow and fast solution processes at time $k\delta$, respectively for $t \in \big[k\delta, \min\{(k+1)\delta, T\}\big), k > 0$. Define the $\hat{X}_t^{\epsilon}$ processes by integral

$$\hat{X}_t^{\epsilon} = x_0 + \int_0^t a(X_{\lfloor s/\delta \rfloor \delta}^{\epsilon}, \hat{Y}_s^{\epsilon})ds + \int_0^t b(X_s^{\epsilon})dB_s^1 + \int_0^t \int_{\mathbb{Z}} \sigma(X_s^{\epsilon}, z)\tilde{N}_1(ds, dz),$$

for $t \in [0, T]$, where $s(\delta) = \lfloor \frac{s}{\delta} \rfloor \delta$ is the nearest breakpoint proceeding s.

Remark 3.2. Using Itô's formula, Assumption (A1) and Lemma 2.4, we can get the corresponding result: $\mathbb{E}\|Y_t^{\epsilon} - \hat{Y}_t^{\epsilon}\|^{2p} \leq C_{p,\iota_i,c_j,T} \frac{\delta^2}{\epsilon} e^{C'_{p,\iota_i,c_j,T} \frac{\delta}{\epsilon}}$. Here, we omit the proof.

Next, it follows from definitions of X_t^{ϵ} and $\hat{X}_t^{\epsilon}$, by Itô's formula, we get

$$\begin{aligned}
&\mathbb{E} \sup_{0 \leq s \leq t} \|X_s^{\epsilon} - \hat{X}_s^{\epsilon}\|^{2p} \\
= \;& 2p\mathbb{E} \sup_{0 \leq s \leq t} \int_0^s \|X_r^{\epsilon} - \hat{X}_r^{\epsilon}\|^{2p-1} \|a\left(X_r^{\epsilon}, Y_r^{\epsilon}\right) - a(X_{\lfloor r/\delta \rfloor \delta}^{\epsilon}, \hat{Y}_r^{\epsilon})\| dr \\
\leq \;& 2p\mathbb{E} \int_0^t \|X_s^{\epsilon} - \hat{X}_s^{\epsilon}\|^{2p-1} \|a\left(X_s^{\epsilon}, Y_s^{\epsilon}\right) - a(X_{\lfloor s/\delta \rfloor \delta}^{\epsilon}, \hat{Y}_s^{\epsilon})\| ds \\
\leq \;& C_p \int_0^t \mathbb{E}\|X_s^{\epsilon} - \hat{X}_s^{\epsilon}\|^{2p} ds + C_p \int_0^t \mathbb{E}\big\|a(X_s^{\epsilon}, Y_s^{\epsilon}) - a(X_{\lfloor s/\delta \rfloor \delta}^{\epsilon}, \hat{Y}_s^{\epsilon})\big\|^{2p} ds \qquad (3.2) \\
\leq \;& C_p \int_0^t \mathbb{E} \sup_{0 \leq r \leq s} \|X_r^{\epsilon} - \hat{X}_r^{\epsilon}\|^{2p} ds + C_p \int_0^t \kappa\big(\mathbb{E}\big\|X_s^{\epsilon} - X_{\lfloor s/\delta \rfloor \delta}^{\epsilon}\big\|^{2p} + \mathbb{E}\big\|Y_s^{\epsilon} - \hat{Y}_s^{\epsilon}\big\|^{2p}\big) ds \\
\leq \;& C_p \int_0^t \mathbb{E} \sup_{0 \leq r \leq s} \|X_r^{\epsilon} - \hat{X}_r^{\epsilon}\|^{2p} ds + C_p\kappa\big(\delta + C_{p,\iota_i,c_j,T} \frac{\delta^2}{\epsilon} e^{C'_{p,\iota_i,c_j,T} \frac{\delta}{\epsilon}}\big) \\
\leq \;& C_p\kappa\big(\delta + C_{p,\iota_i,c_j,T} \frac{\delta^2}{\epsilon} e^{C'_{p,\iota_i,c_j,T} \frac{\delta}{\epsilon}}\big) e^{C_p t} := \varepsilon_0.
\end{aligned}$$

The last term applies the Gronwall's inequality.

To proceed, we can conclude that the mapping $\bar{a} : \mathbb{R}^n \to \mathbb{R}^n$ satisfies non-Lipschitz condition from equation (2.1). Actually, $Y_t^{x_1,y}$ and $Y_t^{x_2,y}$ are two solutions of (2.2), then for any $x_1, x_2 \in \mathbb{R}^n, y \in \mathbb{R}^m$, according to the Kunita's first inequality, one gets

$$\begin{aligned}
\mathbb{E} \sup_{0 \leq s \leq t} \|Y_s^{x_1,y} - Y_s^{x_2,y}\|^{2p} \leq \;& C_p\mathbb{E} \sup_{0 \leq s \leq t} \left\| \int_0^s \big[f(x_1, Y_r^{x_1,y}) - f(x_2, Y_r^{x_2,y})\big] dr \right\|^{2p} \\
& + C_p\mathbb{E} \sup_{0 \leq s \leq t} \left\| \int_0^s \big[g(x_1, Y_r^{x_1,y}) - g(x_2, Y_r^{x_2,y})\big] dB_r^2 \right\|^{2p} \\
& + C_p\mathbb{E} \sup_{0 \leq s \leq t} \left\| \int_0^s \int_{\mathbb{Z}} \big[h(x_1, Y_r^{x_1,y}, z) - h(x_2, Y_r^{x_2,y}), z\big] \tilde{N}_2(dr, dz) \right\|^{2p} \\
\leq \;& C_{p,T}\mathbb{E} \int_0^t \left(\|x_1 - x_2\|^{2p} + \|Y_s^{x_1,y} - Y_s^{x_2,y}\|^{2p}\right) ds \\
\leq \;& C_{p,T}\|x_1 - x_2\|^{2p} + C_{p,T} \int_0^t \mathbb{E} \sup_{0 \leq r \leq s} \|Y_r^{x_1,y} - Y_r^{x_2,y}\|^{2p} ds
\end{aligned}$$

$$\leq \quad C_{p,T}\|x_1 - x_2\|^{2p},$$

where $C_{p,T}$ is independent of (x_1, x_2, y). Moreover, using the property of the function $k(\cdot)$ it is easy to get

$$\begin{aligned}
\|\bar{a}(x_1) - \bar{a}(x_2)\|^{2p} &= \left\|\int_{\mathbb{R}^m} a(x_1, y)\mu^{x_1}(dy) - \int_{\mathbb{R}^m} a(x_2, y)\mu^{x_2}(dy)\right\|^{2p} \\
&\leq \mathbb{E}\,\|a(x_1, Y_t^{x_1,y}) - a(x_2, Y_t^{x_2,y})\|^{2p} \\
&\leq \kappa(\|x_1 - x_2\|^{2p}) + \kappa\left(\mathbb{E}\|Y_t^{x_1,y} - Y_t^{x_2,y}\|^{2p}\right) \\
&\leq C_{p,T}\kappa\left(\|x_1 - x_2\|^{2p}\right).
\end{aligned}$$

This implies that the mapping $\mathbb{R}^n \mapsto \mathbb{R}^n$ is non-Lipchitz continuous. By the linear growth condition of $\bar{a}$, we can derive that the existence and uniqueness of solution of equation (2.1).

Step 2. In this step, we will estimate the following estimate $\mathbb{E}\sup\limits_{0\leq t\leq T}\|\hat{X}_t^\epsilon - \bar{X}_t\|^{2p}$. It follows from the definitions of $\bar{X}_t$ and $\hat{X}_t^\epsilon$ that

$$\begin{aligned}
&\mathbb{E}\sup_{0\leq t\leq T}\|\hat{X}_t^\epsilon - \bar{X}_t\|^{2p} \\
= \;&\mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t \left(a\left(X_{\lfloor s/\delta\rfloor\delta}^\epsilon, \hat{Y}_s^\epsilon\right) - \bar{a}\left(X_s^\epsilon\right)\right)ds\|^{2p} \\
&+\mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t \left(\bar{a}\left(X_s^\epsilon\right) - \bar{a}\left(\hat{X}_s^\epsilon\right)\right)ds\|^{2p} + \mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t \left(\bar{a}\left(\hat{X}_s^\epsilon\right) - \bar{a}\left(\bar{X}_s\right)\right)ds\|^{2p} \\
&+\mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t \left(b\left(X_s^\epsilon\right) - b\left(\hat{X}_s^\epsilon\right)\right)dB_s^1\|^{2p} + \mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t \left(b\left(\hat{X}_s^\epsilon\right) - b\left(\bar{X}_s\right)\right)dB_s^1\|^{2p} \\
&+\mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t\int_{\mathbb{Z}} \left(\sigma\left(X_s^\epsilon, z\right) - \sigma\left(\hat{X}_s^\epsilon, z\right)\right)\tilde{N}_1\left(ds, dz\right)\|^{2p} \\
&+\mathbb{E}\sup_{0\leq t\leq T}\|\int_0^t\int_{\mathbb{Z}} \left(\sigma\left(\hat{X}_s^\epsilon, z\right) - \sigma\left(\bar{X}_s, z\right)\right)\tilde{N}_1\left(ds, dz\right)\|^{2p} := \sum_{i=1}^7 \Theta_i(t).
\end{aligned}$$

Using Hölder inequality, under the help of Burkholder-Davious-Gundy inequality, one obtains

$$\begin{aligned}
\mathbb{E}\Big(\sum_{i=2}^7 \sup_{0\leq t\leq T}\|\Theta_i(t)\|^{2p}\Big) \;\leq\; & C_{p,T}\int_0^T \kappa\Big(\mathbb{E}\sup_{0\leq s\leq t}\|X_s^\epsilon - \hat{X}_s^\epsilon\|^{2p}\Big)dt \\
&+C_{p,T}\int_0^T \kappa\Big(\mathbb{E}\sup_{0\leq s\leq t}\|\bar{X}_s - \hat{X}_s^\epsilon\|^{2p}\Big)dt. \qquad (3.3)
\end{aligned}$$

Now to deal with $\Theta_1(t)$, such that for $t \in [k_t\delta, (k_t+1)\delta\wedge T]$, $k_t := \lfloor \frac{t}{\delta}\rfloor$, we have representation in the form of

$$\Theta_1(t) \;=\; \sum_{i=0}^{k_t-1}\int_{i\delta}^{(i+1)\delta} \left(a(X_{i\delta}^\epsilon, \hat{Y}_s^\epsilon) - \bar{a}(X_{i\delta}^\epsilon)\right)ds + \sum_{i=0}^{k_t-1}\int_{i\delta}^{(i+1)\delta} \left(\bar{a}(X_{i\delta}^\epsilon) - \bar{a}(X_s^\epsilon)\right)ds$$

$$+\int_{k_t\delta}^{t}\big(a(X^\epsilon_{k_t\delta},\hat{Y}^\epsilon_s)-\bar{a}(X^\epsilon_s)\big)ds:=\sum_{i=1}^{3}\Theta_{1i}(t). \tag{3.4}$$

For $\Theta_{12}(t)$, utilizing Hölder inequality and non-Lipschitz continuity of $\bar{a}$, it follows that

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}\|\Theta_{12}(t)\|^{2p} &\le \mathbb{E}\sup_{0\le t\le T}\int_0^{k_t\delta}\big\|\bar{a}(X_{\lfloor\frac{s}{\delta}\rfloor\delta})-\bar{a}(X^\epsilon_s)\big\|^{2p}ds\\ &\le C_{p,T}\int_0^T\kappa\big(\mathbb{E}\|X_{\lfloor\frac{s}{\delta}\rfloor\delta}-X^\epsilon_s\|^{2p}\big)ds. \end{aligned}\tag{3.5}$$

We proceed next to the estimation of $\Theta_{13}(t)$, by linear growth bound of functions a and $\bar{a}$,

$$\mathbb{E}\sup_{0\le t\le T}\|\Theta_{13}(t)\|^{2p}\le\mathbb{E}\sup_{0\le t\le T}\int_{k_t\delta}^{t}\big\|a(X^\epsilon_{k_t\delta},\hat{Y}^\epsilon_s)-\bar{a}(X^\epsilon_s)\big\|^{2p}ds\le C_{p,\iota_i,c_j,T}\delta. \tag{3.6}$$

Again, as for $\Theta_{11}(t)$ we can deduce that

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}\|\Theta_{11}(t)\|^{2p} &= \mathbb{E}\sup_{0\le t\le T}\left\|\sum_{i=0}^{k_t-1}\int_{i\delta}^{(i+1)\delta}\Big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_s)-\bar{a}(X^\epsilon_{i\delta})\Big]ds\right\|^{2p}\\
&\le \mathbb{E}\sup_{0\le t\le T}\big(\frac{t}{\delta}\big)^{2p-1}\sum_{i=0}^{k_t-1}\left\|\int_{i\delta}^{(i+1)\delta}\Big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_s)-\bar{a}(X^\epsilon_{i\delta})\Big]ds\right\|^{2p}\\
&\le \big(\frac{T}{\delta}\big)^{2p-1}\sum_{i=0}^{\lfloor\frac{T}{\delta}\rfloor-1}\mathbb{E}\left\|\int_{i\delta}^{(i+1)\delta}\Big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_s)-\bar{a}(X^\epsilon_{i\delta})\Big]ds\right\|^{2p}\\
&\le \big(\frac{T}{\delta}\big)^{2p}\epsilon^{2p}\max_{0\le i\le\lfloor\frac{T}{\delta}\rfloor-1}\mathbb{E}\left\|\int_0^{\delta/\epsilon}\Big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_{s\epsilon+i\delta})-\bar{a}(X^\epsilon_{i\delta})\Big]ds\right\|^{2p}\\
&\le \big(\frac{\epsilon}{\delta}\big)^{2p}\big(\frac{\delta}{\epsilon}\big)^{2p-2}\max_{0\le i\le\lfloor\frac{T}{\delta}\rfloor-1}\mathbb{E}\left\|\int_0^{\delta/\epsilon}\Big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_{s\epsilon+i\delta})-\bar{a}(X^\epsilon_{i\delta})\Big]ds\right\|^{2}\\
&\le \big(\frac{\epsilon}{\delta}\big)^{2}\max_{0\le i\le\lfloor\frac{T}{\delta}\rfloor-1}\mathcal{I}^\epsilon_k\le C\frac{\epsilon}{\delta},\end{aligned}\tag{3.7}$$

where $\mathcal{I}^\epsilon_i=\mathbb{E}\left\|\int_0^{\delta/\epsilon}\big[a(X^\epsilon_{i\delta},\hat{Y}^\epsilon_{s\epsilon+i\delta})-\bar{a}(X^\epsilon_{i\delta})\big]ds\right\|^2\le C\frac{\delta}{\epsilon}$ (see [13], Lemma 4.1.). Thus, substituting (3.5) (3.6) and (3.7) into (3.4), we can estimate

$$\mathbb{E}\sup_{0\le t\le T}\|\Theta_1(t)\|^{2p}\le C_{p,\iota_i,T}(\delta+\frac{\epsilon}{\delta})+\int_0^T\kappa\big(\mathbb{E}\|X^\epsilon_{\lfloor\frac{s}{\delta}\rfloor\delta}-X^\epsilon_s\|^{2p}\big)ds. \tag{3.8}$$

Taking (3.3) and (3.8) into account, it can be deduce that

$$\mathbb{E}\sup_{0\le t\le T}\|\bar{X}_t-\hat{X}^\epsilon_t\|^{2p}\ \le\ C_{p,T}\int_0^T\kappa\big(\mathbb{E}\sup_{0\le s\le t}\|\bar{X}_s-\hat{X}^\epsilon_s\|^{2p}\big)dt+C_{p,\iota_i,T}(\delta+\frac{\epsilon}{\delta})$$

$$
\begin{aligned}
&+\int_0^T \kappa\big(\mathbb{E}\sup_{0\le s\le t}\|X_s^\epsilon-\hat{X}_s^\epsilon\|^{2p}\big)dt+\int_0^T \kappa\big(\mathbb{E}\|X_{\lfloor \frac{t}{\delta}\rfloor\delta}-X_t\|^{2p}\big)dt\\
\le\ & C_{p,T}\int_0^T \Big[\kappa\big(\mathbb{E}\sup_{0\le s\le t}\|\bar{X}_s-\hat{X}_s^\epsilon\|^{2p}\big)+\mathbb{E}\sup_{0\le s\le t}\|\bar{X}_s-\hat{X}_s^\epsilon\|^{2p}\Big]dt\\
&+C_{p,\iota_i,T}\big(\kappa(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)+(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)\big)\\
\le\ & C_{p,T}\int_0^T \tilde{\kappa}\big(\mathbb{E}\sup_{0\le s\le t}\|\bar{X}_s-\hat{X}_s^\epsilon\|^{2p}\big)dt+C_{p,\iota_i,T}\tilde{\kappa}\big(\delta+\frac{\epsilon}{\delta}+\varepsilon_0\big),
\end{aligned}
$$

where $\tilde{\kappa}(u)=u+\kappa(u)$, obviously, $\tilde{\kappa}(0)=0$, $\int_{0+}\frac{du}{\tilde{\kappa}(u)}=\infty$. Applying the Bihari's inequality to get,

$$
\mathbb{E}\sup_{0\le t\le T}\|\bar{X}_t-\hat{X}_t^\epsilon\|^{2p}\le G^{-1}\Big[G\big(C_{p,\iota_i,T}\tilde{\kappa}(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)\big)+C_{p,T}t\Big],
$$

where $\delta=\epsilon ln\epsilon^{-k}$ $(0<k<\frac{1}{2+C'_{p,\iota_i,c_j,T}})$, $\lim_{\epsilon\to 0}\varepsilon_0=\lim_{\epsilon\to 0}\kappa\big(\delta+C_{p,\iota_i,c_j,T}\frac{\delta^2}{\epsilon}e^{C'_{p,\iota_i,c_j,T}\frac{\delta}{\epsilon}}\big)e^{C_pt}=0$, $\tilde{\kappa}(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)\to 0$, recalling the condition $\int_{0+}\frac{du}{\tilde{\kappa}(u)}=\infty$, we have $G\left(C_{p,\iota_i,T}\tilde{\kappa}(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)\right)+C_{p,T}t\to-\infty$, so we get $G^{-1}\left[G\big(C_{p,\iota_i,T}\tilde{\kappa}(\delta+\frac{\epsilon}{\delta}+\varepsilon_0)\big)+C_{p,T}t\right]\to 0$. Therefore we have

$$
\lim_{\epsilon\to 0}\mathbb{E}\sup_{0\le t\le T}\|\hat{X}_t^\epsilon-\bar{X}_t\|^{2p}=0.
$$

Step 3. In terms of the conclusions of Step 1 and Step 2, we obtain

$$
\begin{aligned}
\mathbb{E}\sup_{0\le t\le T}\|X_t^\epsilon-\bar{X}_t\|^{2p} &= \mathbb{E}\sup_{0\le t\le T}\|X_t^\epsilon-\hat{X}_t^\epsilon+\hat{X}_t^\epsilon-\bar{X}_t\|^{2p}\\
&\le C_p\mathbb{E}\sup_{0\le t\le T}\|X_t^\epsilon-\hat{X}_t^\epsilon\|^{2p}+C_p\mathbb{E}\sup_{0\le t\le T}\|\hat{X}_t^\epsilon-\bar{X}_t\|^{2p}\\
&\le C_p\kappa\big(\delta+C_{p,\iota_i,c_j,T}\frac{\delta^2}{\epsilon}e^{C'_{p,\iota_i,c_j,T}\frac{\delta}{\epsilon}}\big)e^{C_pt}.
\end{aligned}
$$

Consequently, taking $\delta=\epsilon ln\epsilon^{-k}$, letting $\epsilon\to 0$, it is easy to see that Theorem 3.1 holds. The proof of Theorem 3.1 is completed. □

Acknowledgments

This paper was supported by the National Natural Science Foundation of China under Grant No.11772255, the Fundamental Research Funds for the Central Universities, the Research Funds for Interdisciplinary Subject of Northwestern Polytechnical University, the Shaanxi Project for Distinguished Young Scholars, and Shaanxi Provincial Key R&D Program 2020KW-013 and 2019TD-010.

References

References

[1] F.-B. Xi, C. Zhu, Jump type stochastic differential equations with non-Lipschitz coefficients: non confluence, feller and strong feller properties, and exponential ergodicity, Journal of Differential Equations 266 (2019) 4668-4711.

[2] R. Z. Khasminskii, The averaging principle for stochastic differential equations, Problemy Peredachi Informatsii 4 (2) (1968) 86-87.

[3] S. Cerrai, M. Freidlin, Averaging principle for a class of stochastic reaction-diffusion equations, Probability Theory and Related Fields 144 (1-2) (2009) 137-177.

[4] S. Cerrai, A. Lunardi, Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: the almost periodic case, SIAM Journal on Mathematical Analysis 49 (4) (2017) 2843-2884.

[5] D. Givon, Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems, Multiscale Modeling & Simulation 6 (2) (2007) 577-594.

[6] Y. Xu, J.-Q. Duan, W. Xu, An averaging principle for stochastic dynamical systems with Lévy noise, Physica D: Nonlinear Phenomena 240 (17) (2011) 1395-1401.

[7] Y. Xu, B. Pei, Y.-G. Li, Approximation properties for solutions to non-Lipschitz stochastic differential equations with Lévy noise, Mathematical Methods in the Applied Sciences 38 (11) (2015) 2120-2131.

[8] Y. Xu, B. Pei, J.-L. Wu, Stochastic averaging principle for differential equations with non-Lipschitz coefficients driven by fractional Brownian motion, Stochastics and Dynamics 17 (02) (2017) 1750013.

[9] S. Albeverio, Z. Brzeźniak, J.-L. Wu, Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients, Journal of Mathematical Analysis and Applications 371 (1) (2010) 309-322.

[10] J. Xu, J.-C. Liu, Stochastic averaging principle for two-time-scale jump-diffusion SDEs under the non-Lipschitz coefficients, Stochastics-An International Journal of Probability and Stochastic Processes (2020) https://doi.org/10.1080/17442508.2020.1784897.

[11] Y. Guan, J. Wu, Exponential ergodicity for non-Lipschitz multivalued stochastic differential equations with Lévy jumps, Infinite Dimensional Analysis, Quantum Probability and Related Topics 20 (01) (2017) 1750002.

[12] H.-B. Fu, J.-C. Liu, Strong convergence in stochastic averaging principle for two time-scales stochastic partial differential equations, Journal of Mathematical Analysis and Applications 384 (1) (2011) 70-86.

[13] J. Xu, Y. Miao, $L^p (p > 2)$-strong convergence of an averaging principle for two-time-scales jump-diffusion stochastic differential equations, Nonlinear Analysis: Hybrid Systems 18 (2015) 33-47.